Example 2: Probabilistic Learning

2a. Explain how conditional probability works and how it is different from “normal” probability.

2b. Bayesian Networks are shown to account for most conditional probabilities in a scenario. To do so, it uses some variation of the Naïve Bayes Theorem. What is one large caveat that Naïve Bayes has that could prove problematic in producing realistic results?

2c. A company is attempting to create a learning model that produces recommendations for products and services. They currently have the percentage of purchases that consumers have and are attempting to predict what kind of purchases they would make next. To do this, they would like to create a Bayesian network. Would there be any problems in using probabilistic learning for this? For example, if a consumer likes to buy pickles and likes to buy ice cream, does this mean the network will recommend pickle ice cream?

2d. Now say the company above would like to use another Bayesian network to autofill in text if the consumer does not know what they want. For instance, if the user does not know the exact name of an item, they can use previous purchases to tell the Bayesian network. What are some potential problems with using the Bayesian network to perform this task?

2e. Write Bayes Theorem and explain what each part of the equation means in context.

2f. Directed Acyclic Graphs (DAGs) are shown to be an extremely powerful representation in showing the progression of data and flows of control. Suppose a network needs to be built that models conditional probabilities of prior events. However, the initial event will occur once more if the probability of the event prior is high enough. This will make the DAG contain a cycle and will no longer be a true network in the traditional sense. How can this be mitigated using the propagational abilities of the network?